

0590  
06/9

127



OIKE

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/081,806

TIME: 14:00:32

Input Set : N:\Crf3\RULE60\10081806.raw

Output Set: N:\CRF3\07022002\J081806.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Prudent, James R.

6 Hall, Jeff G.

7 Lyamichev, Victor I.

9 (ii) TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids

11 (iii) NUMBER OF SEQUENCES: 69

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Medlen &amp; Carroll, LLP

15 (B) STREET: 220 Montgomery Street, Suite 2200

16 (C) CITY: San Francisco

17 (D) STATE: California

18 (E) COUNTRY: United States Of America

19 (F) ZIP: 94104

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

27 (vi) CURRENT APPLICATION DATA:

C--&gt; 28 (A) APPLICATION NUMBER: US/10/081,806

C--&gt; 29 (B) FILING DATE: 22-Feb-2002

30 (C) CLASSIFICATION:

40 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: US/08/756,386

34 (B) FILING DATE:

37 (A) APPLICATION NUMBER: US 08/682,853

38 (B) FILING DATE: 12-JUL-1996

41 (A) APPLICATION NUMBER: US 08/599,491

42 (B) FILING DATE: 24-JAN-1996

44 (viii) ATTORNEY/AGENT INFORMATION:

45 (A) NAME: Ingolia, Diane E.

46 (B) REGISTRATION NUMBER: 40,027

47 (C) REFERENCE/DOCKET NUMBER: FORS-02564

49 (ix) TELECOMMUNICATION INFORMATION:

50 (A) TELEPHONE: (415) 705-8410

51 (B) TELEFAX: (415) 397-8338

54 (2) INFORMATION FOR SEQ ID NO: 1:

56 (i) SEQUENCE CHARACTERISTICS:

57 (A) LENGTH: 2506 base pairs

58 (B) TYPE: nucleic acid

59 (C) STRANDEDNESS: double

60 (D) TOPOLOGY: linear

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/081,806

DATE: 07/02/2002

TIME: 14:00:32

Input Set : N:\Crf3\RULE60\10081806.raw

Output Set: N:\CRF3\07022002\J081806.raw

```

62      (ii) MOLECULE TYPE: DNA (genomic)
66      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
68 ATGAGGGGGA TGCTGCCCCCT CTTTGAGCCC AAGGGCCGGG TCCTCCTGGT GGACGGCCAC      60
70 CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG GGGGGAGCCG      120
72 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCTCAAGG CCCTCAAGGA GGACGGGGAC      180
74 GCGGTGATCG TGGTCTTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC CTACGGGGGG      240
76 TACAAGGCGG GCCGGGCCCC CACGCCGGAG GACTTTCCCC GGCAACTCGC CCTCATCAAG      300
78 GAGCTGGTGG ACCTCCTGGG GCTGGCGCGC CTCGAGGTCC CGGGCTACGA GGCGGACGAC      360
80 GTCCTGGCCA GCCTGGCCAA GAAGGCGGAA AAGGAGGGCT ACGAGGTCCG CATCCTCACC      420
82 GCCGACAAAG ACCTTTACCA GCTCCTTTCC GACCGCATCC ACGTCCTCCA CCCCAGGGGG      480
84 TACCTCATCA CCCCAGCCTG GCTTTGGGAA AAGTACGGCC TGAGGCCCCG CCAGTGGGCC      540
86 GACTACCGGG CCGTGACCGG GGACGAGTCC GACAACCTTC CCGGGGTCAA GGGCATCGGG      600
88 GAGAAGACGG CGAGGAAGCT TCTGGAGGAG TGGGGGAGCC TGGGAAGCCCT CCTCAAGAAC      660
90 CTGGACCGGC TGAAGCCCGC CATCCGGGAG AAGATCCTGG CCCACATGGA CGATCTGAAG      720
92 CTCTCCTGGG ACCTGGCCAA GGTGCGCACC GACCTGCCCC TGGAGGTGGA CTTGCGCCAA      780
94 AGGCGGGAGC CCGACCGGGA GAGGCTTAGG GCCTTTCTGG AGAGGCTTGA GTTTGGCAGC      840
96 CTCTCCACG AGTTCGGCCT TCTGGAAAGC CCAAGGCCC TGGAGGAGGC CCCCTGGCCC      900
98 CCGCCGAAG GGGCCTTCGT GGGCTTTGTG CTTTCCCGCA AGGAGCCCAT GTGGGCCGAT      960
100 CTCTGGCCC TGGCCGCCGC CAGGGGGGGC CGGCTCCACC GGGCCCCCGA GCCTTATAAA      1020
102 GCCCTCAGGG ACCTGAAGGA GGCGCGGGGG CTTCTCGCCA AAGACCTGAG CGTTCTGGCC      1080
104 CTGAGGGAAG GCCTTGGCCT CCCGCCGGC GACGACCCCA TGCTCCTCGC CTACCTCCTG      1140
106 GACCTTCCA ACACCACCC CGAGGGGGTG GCCCGGCGCT ACGGCGGGGA GTGGACGGAG      1200
108 GAGGCGGGGG AGCGGGCCGC CTTTCCGAG AGGCTCTTCG CCAACCTGTG GGGGAGGCTT      1260
110 GAGGGGGAGG AGAGGCTCCT TTGGCTTTAC CGGGAGGTGG AGAGGCCCTT TTCCGCTGTC      1320
112 CTGGCCCACA TGGAGCCAC GGGGTGCGC CTGGACGTGG CCTATCTCAG GGCCTTGTCT      1380
114 CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC GAGGCCGAGG TCTTCCGCCT GGCCGGCCAC      1440
116 CCCTTCAACC TCAACTCCCG GGACCAGCTG GAAAGGTCC TCTTTGACGA GCTAGGGCTT      1500
118 CCCGCCATCG GCAAGACGGA GAAGACCGGC AAGCGTCCA CCAGCGCCGC CGTCCTGGAG      1560
120 GCCCTCCGCG AGGCCACCC CATCGTGGAG AAGATCCTGC AGTACCGGGA GCTACCAAG      1620
122 CTGAAGAGCA CCTACATTGA CCCCTTGCCG GACCTCATCC ACCCCAGGAC GGGCCGCCTC      1680
124 CACACCCGCT TCAACCAGAC GGCCACGGCC ACGGGCAGGC TAAGTAGCTC CGATCCCAAC      1740
126 CTCCAGAACA TCCCCGTCCG CACCCGCTT GGGCAGAGGA TCCGCCGGGC CTTTCATCGCC      1800
128 GAGGAGGGGT GGCTATTGGT GGCCCTGGAC TATAGCCAGA TAGAGCTCAG GGTGCTGGCC      1860
130 CACCTCTCCG GCGACGAGAA CCTGATCCGG GTCTTCCAGG AGGGGCGGGA CATCCACACG      1920
132 GAGACCGCCA GCTGGATGTT CGGCGTCCCC CGGGAGGCCG TGGACCCCTT GATGCGCCGG      1980
134 GCGGCCAAGA CCATCAACTT CGGGTTCCTC TACGGCATGT CGGCCACCG CCTCTCCAG      2040
136 GAGCTAGCCA TCCCTTACGA GGAGGCCAG GCCTTCATTG AGCGCTACTT TCAGAGCTTC      2100
138 CCAAGGTGC GGGCCTGGAT TGAGAAGACC CTGGAGGAGG GCAGGAGGCG GGGGTACGTG      2160
140 GAGACCTCT TCGGCCGCG CCGCTACGTG CCAGACCTAG AGGCCCGGGT GAAGAGCGTG      2220
142 CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC ATGCCCGTCC AGGGCACCGC CGCCGACCTC      2280
144 ATGAAGCTGG CTATGGTGAA GCTCTTCCCC AGGCTGGAGG AAATGGGGGC CAGGATGCTC      2340
146 CTTCAGGTCC ACGACGAGCT GGTCTCGAG GCCCCAAAAG AGAGGGCGGA GGCCGTGGCC      2400
148 CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG TATCCCTGG CCGTGCCCCT GGAGGTGGAG      2460
150 GTGGGGATAG GGGAGGACTG GCTCTCCGCC AAGGAGTGAT ACCACC      2506
152 (2) INFORMATION FOR SEQ ID NO: 2:
154      (i) SEQUENCE CHARACTERISTICS:
155          (A) LENGTH: 2496 base pairs
156          (B) TYPE: nucleic acid
157          (C) STRANDEDNESS: double

```

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/081,806

TIME: 14:00:32

Input Set : N:\Crif3\RULE60\10081806.raw

Output Set: N:\CRF3\07022002\J081806.raw

```

158      (D) TOPOLOGY: linear
160      (ii) MOLECULE TYPE: DNA (genomic)
164      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
166 ATGGCGATGC TTCCCTCTT TGAGCCCAAA GGCCGCGTGC TCCTGGTGGA CGGCCACCAC      60
168 CTGGCCTACC GCACCTTCTT TGCCCTCAAG GGCCTCACCA CCAGCCGCGG CGAACCCGTT      120
170 CAGGCGGTCT ACGGCTTCGC CAAAAGCCTC CTCAAGGCCC TGAAGGAGGA CGGGGACGTG      180
172 GTGGTGGTGG TCTTTGACGC CAAGGCCCCC TCCTTCCGCC ACGAGGCCTA CGAGGCCTAC      240
174 AAGGCGGGCC GGGCCCCCAC CCCGGAGGAC TTTCCCCGGC AGCTGGCCCT CATCAAGGAG      300
176 TTGGTGGACC TCCTAGGCCT TGTGCGGCTG GAGGTTCCCG GCTTTGAGGC GGACGACGTG      360
178 CTGGCCACCC TGGCCAAGCG GCGGAAAAG GAGGGGTACG AGGTGCGCAT CCTCACTGCC      420
180 GACCGCGACC TCTACGAGCT CCTTTCGGAG CGCATCGCCA TCCTCCACCC TGAGGGGTAC      480
182 CTGATCACCC CGGCGTGGCT TTACGAGAAG TACGGCCTGC GCCCGAGCA GTGGGTGGAC      540
184 TACCGGGCCC TGGCGGGGGA CCCCTCGGAT AACATCCCCG GGGTGAAGGG CATCGGGGAG      600
186 AAGACCGCCC AGAGGCTCAT CCGCGAGTGG GGGAGCCTGG AAAACCTCTT CCAGCACCTG      660
188 GACCAGGTGA AGCCCTCCTT GCGGGAGAAG CTCCAGGCGG GCATGGAGGC CCTGGCCCTT      720
190 TCCCGGAAGC TTTCCAGGT GCACACTGAC CTGCCCTGG AGGTGGACTT CGGGAGGCGC      780
192 CGCACACCCA ACCTGGAGGG TCTGCGGCT TTTTGGAGC GGTGGAGTT TGGAAGCCTC      840
194 CTCCACGAGT TCGGCCTCCT GGAGGGGCGG AAGGCGGCAG AGGAGGCCCC CTGGCCCCCT      900
196 CCGGAAGGGG CTTTTTTGGG CTTTTCTTTT TCCCGTCCCG AGCCCATGTG GGCCGAGCTT      960
198 CTGGCCCTGG CTGGGGCGTG GGAGGGGCGC CTCCATCGGG CACAAGACCC CCTTAGGGGC      1020
200 CTGAGGGACC TTAAGGGGGT GCGGGGAATC CTGGCCAAGG ACCTGGCGGT TTTGGCCCTG      1080
202 CGGGAGGGCC TGGACCTCTT CCCAGAGGAC GACCCCATGC TCCTGGCCTA CCTTCTGGAC      1140
204 CCCTCCAACA CCACCCCTGA GGGGGTGGCC CGGCGTTACG GGGGGGAGTG GACGGAGGAT      1200
206 GCGGGGAGAG GGGCCCTCCT GGCCGAGCGC CTCTTCCAGA CCTAAAGGA GCGCCTTAAG      1260
208 GGAGAAGAAC GCCTGCTTTG GCTTTACGAG GAGGTGGAGA AGCCGCTTTC CCGGGTGTTG      1320
210 GCCCGGATGG AGGCCACGGG GGTCCGGCTG GACGTGGCCT ACCTCCAGGC CCTCTGCCCTG      1380
212 GAGGTGAGG CGGAGGTGCG CCAGCTGGAG GAGGAGGTCT TCCGCTGGC CGGCCACCCC      1440
214 TTCAACCTCA ACTCCGCGA CCAGCTGGAG CGGGTGCTCT TTGACGAGCT GGGCCTGCCT      1500
216 GCCATCGGCA AGACGGAGAA GACGGGAAA CGCTCCACCA GCGCTGCCGT GCTGGAGGCC      1560
218 CTGCGAGAGG CCCACCCCAT CGTGGACCGC ATCCTGCAGT ACCGGGAGCT CACCAAGCTC      1620
220 AAGAACACCT ACATAGACCC CCTGCCCGCC CTGGTCCACC CCAAGACCGG CCGGCTCCAC      1680
222 ACCCGCTTCA ACCAGACGGC CACCGCCACG GGCAGGCTTT CCAGCTCCGA CCCCACCTG      1740
224 CAGAACATCC CCGTGCGCAC CCCTCTGGGC CAGCGCATCC GCCGAGCCTT CGTGGCCGAG      1800
226 GAGGGCTGGG TGCTGGTGGT CTTGGACTAC AGCCAGATTG AGCTTCGGGT CCTGGCCAC      1860
228 CTCTCCGGGG ACGAGAACCT GATCCGGGTC TTTCAGGAGG GGAGGGACAT CCACACCCAG      1920
230 ACCGCCAGCT GGATGTTCCG CGTTTCCCCC GAAGGGGTAG ACCCTCTGAT GCGCCGGGCG      1980
232 GCCAAGACCA TCAACTTCGG GGTGCTCTAC GGCATGTCCG CCCACCGCCT CTCCGGGGAG      2040
234 CTTTCCATCC CCTACGAGGA GGCGGTGGCC TTCATTGAGC GCTACTTCCA GAGCTACCCC      2100
236 AAGGTGCGGG CCTGGATTGA GGGGACCTC GAGGAGGGCC GCCGGCGGGG GTATGTGGAG      2160
238 ACCCTCTTCG GCGGCGGCG CTATGTGCC GACCTCAACG CCCGGGTGAA GAGCGTGCGC      2220
240 GAGGCGGCGG AGCGCATGGC CTTCAACATG CCGGTCCAGG GCACCGCCGC CGACCTCATG      2280
242 AAGCTGGCCA TGGTGC GGCT TTTCCCCCGG CTTCAGGAAC TGGGGGCGAG GATGCTTTTG      2340
244 CAGGTGCACG ACGAGCTGGT CCTCGAGGCC CCAAGGACC GGGCGGAGAG GGTAGCCGCT      2400
246 TTGGCCAAGG AGGTCAATGA GGGGTCTGG CCCCTGCAGG TGCCCTGGA GGTGGAGGTG      2460
248 GGCCTGGGGG AGGACTGGCT CTCCGCCAAG GAGTAG      2496
250 (2) INFORMATION FOR SEQ ID NO: 3:
252      (i) SEQUENCE CHARACTERISTICS:
253          (A) LENGTH: 2504 base pairs
254          (B) TYPE: nucleic acid

```

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/081,806

TIME: 14:00:32

Input Set : N:\Crif3\RULE60\10081806.raw

Output Set: N:\CRF3\07022002\J081806.raw

```

255      (C) STRANDEDNESS: double
256      (D) TOPOLOGY: linear
258      (ii) MOLECULE TYPE: DNA (genomic)
262      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
264 ATGGAGGCGA TGCTTCCGCT CTTTGAACCC AAAGGCCCGG TCCTCCTGGT GGACGGCCAC      60
266 CACCTGGCCT ACCGCACCTT CTTCGCCCTG AAGGGCCTCA CCACGAGCCG GGGCGAACCG      120
268 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCTGAAGGA GGACGGGTAC      180
270 AAGGCCGTCT TCGTGGTCTT TGACGCCAAG GCCCCTCCTT TCCGCCACGA GGCTTACGAG      240
272 GCCTACAAGG CGGGGAGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC      300
274 AAGGAGCTGG TGGACCTCCT GGGGTTTACC CGCCTCGAGG TCCCCGGCTA CGAGGCGGAC      360
276 GACGTTCTCG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT GCGCATCCTC      420
278 ACCGCCGACC GCGACCTCTA CCAACTCGTC TCCGACCGCG TCGCCGTCCT CCACCCCGAG      480
280 GGCCACCTCA TCACCCCGGA GTGGCTTTGG GAGAAGTACG GCCTCAGGCC GGAGCAGTGG      540
282 GTGGACTTCC GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC      600
284 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGA AAA CCTCCTCAAG      660
286 AACCTGGACC GGTAAAGCC AGAAAACGTC CGGGAGAAGA TCAAGGCCCA CCTGGAAGAC      720
288 CTCAGGCTCT CTTGGAGCT CTCCCGGGTG CGCACCGACC TCCCCCTGGA GGTGGACCTC      780
290 GCCCAGGGGC GGGAGCCCGA CCGGGAGGGG CTTAGGGCCT TCCTGGAGAG GCTGGAGTTC      840
292 GGCAGCCTCC TCCACGAGTT CGGCCTCCTG GAGGCCCCCG CCCCCCTGGA GGAGGCCCCC      900
294 TGGCCCCCGC CGGAAGGGGC CTTCTGTGGC TTCTCTCTCT CCCGCCCCGA GCCCATGTGG      960
296 GCGGAGCTTA AAGCCCTGGC CGCCTGCAGG GACGGCCGGG TGCACCGGGC AGCAGACCCC      1020
298 TTGGCGGGGC TAAAGGACCT CAAGGAGGTC CGGGGCCTCC TCGCCAAGGA CCTCGCCGTC      1080
300 TTGGCCTCGA GGGAGGGGCT AGACCTCGTG CCCGGGGACG ACCCATGCT CCTCGCCTAC      1140
302 CTCCTGGACC CCTCCAACAC CACCCCGAGG GGGGTGGCGC GCGCTACGG GGGGAGTGG      1200
304 ACGGAGGACG CCGCCACCG GGCCTCCTC TCGGAGAGGC TCCATCGGA CCTCCTTAAG      1260
306 CGCCTCGAGG GGGAGGAGAA GCTCCTTTGG CTCTACCACG AGGTGGAAAA GCCCCTCTCC      1320
308 CGGGTCTTGG CCCACATGGA GGCCACCGGG GTACGGCTGG ACGTGGCCTA CCTTCAGGCC      1380
310 CTTTCCCTGG AGCTTGCGGA GGAGATCCGC CGCCTCGAGG AGGAGGTCTT CCGCTTGGCG      1440
312 GGCCACCCCT TCAACCTCAA CTCCCGGGAC CAGCTGGAAA GGGTGCTCTT TGACGAGCTT      1500
314 AGGCTTCCCG CCTTGGGGAA GACGCAAAAG ACAGGCAAGC GCTCCACCAG CGCCGCGGTG      1560
316 CTGGAGGCCC TACGGGAGGC CCACCCATC GTGGAGAAGA TCCTCCAGCA CCGGGAGCTC      1620
318 ACCAAGCTCA AGAACACCTA CGTGGACCCC CTCCCAAGCC TCGTCCACCC GAGGACGGGC      1680
320 CGCCTCCACA CCCGTTCAA CCAGACGGCC ACGGCCACGG GGAGGCTTAG TAGCTCCGAC      1740
322 CCCAACCTGC AGAACATCCC CGTCCGACCC CCCTTGGGGC AGAGGATCCG CCGGGCCTTC      1800
324 GTGGCCGAGG CGGGTTGGGC GTTGGTGGCC CTGGACTATA GCCAGATAGA GCTCCGCGTC      1860
326 CTCGCCACCC TCTCCGGGGA CGAAAACCTG ATCAGGGTCT TCCAGGAGGG GAAGGACATC      1920
328 CACACCCAGA CCGCAAGCTG GATGTTCCGC GTCCCCCGG AGGCCGTGGA CCCCCTGATG      1980
330 CGCCGGGCGG CCAAGACGGT GAACTTCCGC GTCTCTACG GCATGTCCGC CCATAGGCTC      2040
332 TCCCAGGAGC TTGCCATCCC CTACGAGGAG GCGGTGGCCT TTATAGAGGC TACTTCCAAA      2100
334 GCTTCCCCAA GGTGCGGGCC TGGATAGAAA AGACCTTGA GGAGGGGAGG AAGCGGGGCT      2160
336 ACGTGGAAAC CCTCTTCGGA AGAAGGCGCT ACGTGCCCGA CCTCAACGCC CGGGTGAAGA      2220
338 GCGTCAGGGA GGCCGCGGAG CGCATGGCCT TCAACATGCC CGTCCAGGGC ACCGCCGCCG      2280
340 ACCTCATGAA GCTCGCCATG GTGAAGCTCT TCCCCCGCCT CCGGGAGATG GGGGCCCGCA      2340
342 TGCTCCTCCA GGTCCACGAC GAGCTCCTCC TGGAGGCCCC CCAAGCGCGG GCCGAGGAGG      2400
344 TGGCGGCTTT GGCCAAGGAG GCCATGGAGA AGGCCATATC CCTCGCCGTG CCCCTGGAGG      2460
346 TGGAGGTGGG GATGGGGGAG GACTGGCTTT CCGCCAAGGG TTAG      2504
348 (2) INFORMATION FOR SEQ ID NO: 4:
350      (i) SEQUENCE CHARACTERISTICS:
351      (A) LENGTH: 832 amino acids

```

## RAW SEQUENCE LISTING

DATE: 07/02/2002

PATENT APPLICATION: US/10/081,806

TIME: 14:00:32

Input Set : N:\Crf3\RULE60\10081806.raw

Output Set: N:\CRF3\07022002\J081806.raw

```

352      (B) TYPE: amino acid
353      (C) STRANDEDNESS: single
354      (D) TOPOLOGY: linear
356      (ii) MOLECULE TYPE: protein
360      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
362      Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
363      1          5          10          15
365      Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
366      20          25          30
368      Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
369      35          40          45
371      Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
372      50          55          60
374      Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
375      65          70          75          80
377      Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
378      85          90          95
380      Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
381      100         105         110
383      Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
384      115         120         125
386      Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
387      130         135         140
389      Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
390      145         150         155         160
392      Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
393      165         170         175
395      Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
396      180         185         190
398      Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
399      195         200         205
401      Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
402      210         215         220
404      Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
405      225         230         235         240
407      Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
408      245         250         255
410      Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
411      260         265         270
413      Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
414      275         280         285
416      Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
417      290         295         300
419      Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
420      305         310         315         320
422      Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
423      325         330         335
425      Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
426      340         345         350

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/081,806

DATE: 07/02/2002  
TIME: 14:00:33

Input Set : N:\Crf3\RULE60\10081806.raw  
Output Set: N:\CRF3\07022002\J081806.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 4,5,181,182,190,366,617,628,685,714,722,738,784,1022,1029  
Seq#:7; N Pos. 1038,1053,1098,1105,1206,1227,1244,1251,1252,1253,1350,1380  
Seq#:7; N Pos. 1497,1530,1569,1572,1641,1653,1655,1770,1812,2319,2346,2396  
Seq#:8; Xaa Pos.2,63,109,186,205,209,227,228,233,240,243,244,247,260,290  
Seq#:8; Xaa Pos.329,336,340,368,414,417,418,431,551,605,773,794,798,823,833  
Seq#:57; N Pos. 18

## VERIFICATION SUMMARY

DATE: 07/02/2002

PATENT APPLICATION: US/10/081,806

TIME: 14:00:33

Input Set : N:\Crf3\RULE60\10081806.raw

Output Set: N:\CRF3\07022002\J081806.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:974 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0  
L:983 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:48  
L:992 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:96  
L:1007 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:176  
L:1010 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:192  
L:1013 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:208  
L:1016 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224  
L:1019 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:240  
L:1022 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:256  
L:1028 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:288  
L:1034 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:320  
L:1037 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:336  
L:1040 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:352  
L:1049 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:400  
L:1052 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:416  
L:1076 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:544  
L:1085 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:592  
L:1118 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:768  
L:1121 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:784  
L:1127 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:816  
L:1130 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:832  
L:2269 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=56